

Marek STACHOWSKI (Kraków)

REMARKS ON THE INVESTIGATION OF THE OLDEST LAYER OF TURKIC LOAN-WORDS IN HUNGARIAN

Abstract. Those who are interested in the investigation of the oldest Turkic loan-words into Hungarian have known for a couple of years about the research on this topic conducted in Szeged. It has resulted in a two-volume edition which will certainly inspire many scholars from now onwards. In the present article a handful of remarks and suggestions is presented that were noted down while reading the “new Gombocz”.

Keywords: etymology, Turkic, Hungarian, loan-words, Proto-Altaic

András Róna-Tas / Árpád Bertas, with the assistance of László Károly:
West Old Turkic. Turkic loanwords in Hungarian (= *Turcologica* 84), Wiesbaden
2011, vol. I–II.

1.

A new work on the oldest layer of Turkic loan-words in Hungarian has been a linguistic desideratum for many years. The two-volume study that we now have on our desks is an accurately edited *opus* comprising an informative introduction (I 3–49), the lexicon (I 53–618; II 620–1008), a historical phonology of Hungarian (II 1011–1069), chapters on West Old Turkic and Hungarian (II 1071–1146) and on the question of “Who spoke West Old Turkic, when and where?” (II 1147–1176), as well as indexes, bibliography and so on. This all makes a very good impression. Nevertheless, a reader who is well-read in the interpretations and reconstructions given in this work will sometimes wonder if the term “West Old Turkic” (= WOT) is not just another name for “Proto-Altaic” (indeed, some WOT reconstructions look at times more Altaistic than the authors of the work could enthusiastically accept) and he will certainly feel compelled to pose some additional questions.

Apropos of the $l^1 : l^2$ opposition one can read the following statement here: “This dichotomy was later accepted by many scholars in various ways. This sup-

position is without foundation, since we have no phonological opposition” (II 1105). This is a rather bold opinion. Even G. Doerfer, a much-cited anti-Altaistic author has made this distinction (e.g. in TMEN I 98) even though its rejection would suit him perfectly. The fact that Čuv. *l* has two different equivalents in non-Bulgarian Turkic languages (cf. Čuv. *śula*- = Tkc.¹ *jal(g)a*- ‘to lick’; Čuv. *śul* = Tkc. *jaś* ‘tear’) makes us conjecture the existence of two sources originally different in some way, say *l*¹ and *l*². But the most astonishing thing is the argument that *l*¹ and *l*² could not have possibly existed because “we have no phonological opposition”. Sounds that are allophones in complementary distribution are very likely to develop differently. In short: There is no direct connection between the phonological status of two sounds (phonemes or allophones) on the one hand and their reconstructions based on different modern reflexes on the other.

Čuv. *śāl* ‘tooth’ (= Tkc. *tiś* id.) is cited as a proof for an original *-*ś* that triggered the assimilation of the word-initial **č*- (< **ti*), i.e. **tiś* > **čiš* > **śiš* > Čuv. *śāl* (II 1106). It is, however, rather risky to offer a unique solution on the basis of one example only. We have some Chuvash words with the **ž* > Čuv. *ś* change, e.g. WOT **žalym* > Čuv. *śilēm* ‘a kind of fishing net’ (I 364), WOT **žu-l-* > Čuv. *śu-n-* ‘to burn (intrans.)’. In none of these examples is the second consonant *-*ś* so that assimilation is out of the question. Nevertheless, the WOT **ž*- > Čuv. *ś*- change did occur. Analogically, also the WOT (**ti*- >) **č*- > Čuv. *ś*- change seems to be easily possible. Unfortunately, the authors do not discuss this possibility.

The authors are always in favour of the anti-Altaistic interpretation, that is, they speak of rhotacism and lambdaism only, never of zetacism and sigmatism. They have of course the right to do so. However, at least one factor gives us cause for concern. On page 1107sq. a reader will find a great number of examples for rhotacism and lambdaism, adduced from different languages but not a single example for zetacism and sigmatism is cited. This makes a totally incorrect impression that zetacism and sigmatism are either completely unknown or, at the best, only scarcely encountered in the world’s languages. This is of course not true but it helps to convince readers who may not realise that the existence of rhotacistic and lambdaistic processes in many languages does not prove in itself that Turkic consonants also developed along the same lines.

In addition, the authors have misunderstood one process: It is true that “[i]n Czech the palatalized /ř/ became /ṛ̌/” (II 1107); but the development was zetacism, not rhotacism, which means that this process is incorrectly classified by the authors here, its next evolutionary stage being Polish <ř> = [ž] (< *ř) whose pronunciation today is absolutely identical with that of Polish <ž> = [ž] (< *ž).

The chronology of the rhotacism, too, has been established, in this book, on the basis of one word only: *izengü* ‘stirrup’. Other criteria, arguments and

¹ The abbreviation “Tkc.” stands for “(Standard, i.e. non-Bulgarian) Turkic”.

opinions are omitted altogether. The dating of this process is said by the authors to be as follows: “Since rhotacism occurred in this word [i.e. in *izengü* – M. S.], the phonetic change resulting in Chuv[ash] /r/ and C[ommon]T[urkic] /z/ [...] must date to a period after the stirrup was invented (c3rd c. B.C.) and spread throughout Europe (3th [!], pro: 3rd – M. S.] – 5th c. A.D.) [...]” (II 1112sq.). The formulation is not really clear; the use of the conjunction *and* (in: “was invented [...] and spread”) suggests that the rhotacism should have occurred after the 5th century A.D. (otherwise the dating of the spread of the stirrup is unnecessary in this context). This conclusion would conflict, however, with the earlier opinion on the question given by A. Róna-Tas² thus: “The period when the rotacism [!, pro: rho... – M. S.] has been finished and the lambdaism ceased to work was in the few hundred years before Christ” (Róna-Tas 1999: 13, fn. 18). Despite the vagueness attending the expression *the few hundred years*, it is clear this dating contradicts the idea that the spread of the stirrup is of some importance to us.

As far as other views are concerned I would like to cite, first of all, J. Benzing (1944: 27): “Da wir [...] allen Grund zur Annahme haben, daß der Wandel $z > r$ [...] im Bolgarischen nicht vor dem XI. Jahrh. erfolgt ist, können die Ungarn die bisher für bolgarisch gehaltenen Lehnwörter keinesfalls von den Bulgaren übernommen haben, denn um diese Zeit hatten sie keine Berührung mehr mit den Bulgaren.” I fully agree that Benzing’s opinion is incorrect; nevertheless, where should it be discussed if not in this work? Alas, it is not even mentioned in the chapter “The chronology of rhotacism” (II 1112sq.). Neither is A. M. Ščerbak who in one and the same work first says that the rhotacism and lambdaism are relatively recent phenomena (“[...] рота́цизм и ла́мбдаи́зм в чувашском языке – довольно поздние явления”, Ščerbak 1970: 85) and then, three pages later, that they are developments of considerable antiquity (“Обнаружение следов рота́цизма и ла́мбдаи́зма в общетюркском масштабе является достаточно убедительным доводом его значительной древности”, op.cit. 88). One cannot but regret that these opinions are not commented upon here by such an experienced linguist as A. Róna-Tas.

Worse still, the graphematic criterion is completely omitted here, even if A. Róna-Tas was among first scholars who investigated this aspect of the Old Turkic runes. In short: the rune <z> was invented on the basis of <r>, the rune <š> on the basis of <l>. This fact not only contradicts the anti-Altaistic interpretation that $r < *z$, and $l < *š$, but also suggests that there must have been, regardless of the direction of the sound changes, some sort of phonetic similarity between *r* and *z* on the one hand and *l* and *š* on the other at the time when new runes were

² Although two main authors are mentioned on the title page (in the non-alphabetical order!), A. Róna-Tas fairly regularly uses the 1st person singular pronoun in the text of this book, e.g. “I was able to witness and describe [...] (see Róna-Tas [...])” (II 1162) and much else.

invented. The oldest Turkic runic inscriptions were made in the early 8th century; it is, thus, imaginable that the runes for <z> and <š> were introduced in the late 7th century. This means that the phonetic change has not yet been completed in the 7th century A.D. (for further details see Stachowski 1998) so that neither the rhotacism “has been finished” nor the lambdaism “ceased to work [...] in the few hundred years before Christ” (cf. above: Róna-Tas 1999: 13, fn. 18).

Even if we cannot be sure about the beginning of the zetacism/rhotacism and sigmatism/lambdaism the processes were still in progress in the 5th and 6th century, that is at the time when Proto-Hungarians could have met Turkic peoples. In other words: at that time the **r²* and **l²* pronunciation was typical of *all* the Turkic languages which means that this feature cannot be understood as a signal of the Bulgarian Turkic origin of a word.

Therefore my ultimate conclusion is the same as that of Benzing’s, albeit on quite different counts.

2.

The bulk of this book is imposing, indeed. But it arises, at least to some extent, from long or very long lists of adduced words. Attestations for the verb *ač-* ‘to open’ take up more than two pages (II 1214–1217) and are actually rather unnecessary because the phonetic shape and the meaning of this verb are generally identical in all the Turkic languages. In the remainder of the discussion one learns that the form *ač-* and the meaning ‘to open’ are exactly the data necessary for the discussion of the Hungarian etymology (incidentally, one in the end dismissed by the authors). For another example of this kind see *bögöly* below.

It is absolutely impossible to discuss every etymology in detail here. I have to confine myself to some remarks only:

agár ‘greyhound’ (II 1230). – This hunting term certainly is not of Slavonic origin in Hungarian, contrary to the authors’ opinion. Neither can its Proto-Turkic form be reconstructed with front vowels, e.g. **ägär*. For arguments and discussion see Stachowski M. 2003.

alma ‘apple’ (I 62). – For this extremely complex etymology see now Piwowarczyk 2014 (in this volume of SEC).

ápol ‘to nurse, take care of’ (I 64). – In this context, Tk. *obur* and Slav. *upir* ‘vampire’, too, are rightly mentioned (the word *perhaps* in the formulation “and perhaps Grm [= German] *Vampir*” (I 65) can be readily deleted). However, the authors’ contentment with a single reference to the Russian edition of M. Vasmer’s dictionary (certainly somewhat outdated for this purpose) is a surprise, given the bulk of professional literature on this topic. For *obur* ~

- upir* ~ *vampire* see now Stachowski K. 2005 (esp. the scheme on p. 87 that can be easily understood even without a command of Polish).
- árpa** ‘barley’ (I 77). – Among hitherto suggested etymologies of Tkc. *arpa* id., the only “purely Turkic” one (i.e. that by B. I. Tatarincev), is missing here, see Stachowski K. 2008: 13.
- barom** ‘cattle, livestock’ (I 99sq.). – In this context Slav. *baran* ‘ram, male sheep’, too, is mentioned. One can easily accept the opinion that the Slavonic word is not of Turkic origin as it cannot be compared with any Turkic word without additional explanations of mostly non-trivial differences. Nevertheless, the authors’ view that the Slavonic word is, instead, “more likely of onomatopoeic origin” (I 102) cannot be accepted either (even if this opinion is not really new in the history of etymologizing of this word).³ Its morphological structure is unclear in any Slavonic language and one wonders if **bar* (or the like) is actually an adequate rendering of the sound made by a ram. Besides, the idea of onomatopoeic origin does not allow for possible connections of this word family with Spanish *marrano* ‘male pig’ and Catalanian *marrà* ‘ram, male sheep’. The history of Slavonic *baran* seems to be much more complex than suggested by the authors of this book. It cannot thus be excluded either that Hung. *barom* ‘livestock’, too, belongs to the same word family that appears to be represented in rather distant languages, or that its phonetic resemblance to Turkic both *bār* ‘there is, there exists’ and *bār-* ‘to go’ is merely accidental.
- béka** ‘frog’ (I 110 sq.). – When speaking of the onomatopoeic origin of this word, Á. Berta’s article of 2000 is mentioned. However, this idea was earlier suggested in Li 1997.
- borz** ‘badger, *Meles meles*’ (I 159). – The etymon of this word is reconstructed as WOT (= West Old Turkic) **borsy* = EOT (= East Old Turkic) **borsuk*. This means that these Turkic forms are morphologically different. However, another word, namely Hung. *bíró* ‘judge’ < WOT **byrug* = EOT **byruk*, a title (I 130) shows another possible way: WOT **borsyg* = EOT **borsyk* (> **borsuk*). The WOT **borsyg* would have first produced a spirantic variant **borsyy* and then yielded a form with a long vowel in Hungarian, say, **borsȳ* (cf. II 1081: WOT **bolčuy* > Hung. **bulčuy* > **bulčū* > *bűcű* ‘farewell’). Admittedly, word-final long vowels were usually shortened but not dropped. However, they were in rare cases first shortened and then dropped, as for instance in EOT **butyk* ‘branch of a tree, twig’ = WOT **butyγ* [rather WOT **butyg* > **butyy* – M. S.] ~ **butiy* > **butu* [probably a misprint for **butū* – M. S.] ~ **butī* > **buti* > Hung. *but* (~ *bot*) ‘stick, cane’ (I 161). This being the case, also a parallel change of Hung. **borsȳ* > **borsy* > **bors* is thinkable.

³ Certain doubts concerning the possibility of explaining the origin of this word are presented in Stachowski M. (2005: 438–441).

bögöly ‘horsefly, gadfly’ (I 167). – The authors of this book are, as mentioned above, very generous with space for enumerating Turkic attestations, even those of no comparative importance. Three pages are used up for different attestations of *tüş-* ‘to fall down, settle down’, together with whole sentences⁴, idiomatic expressions⁵ and modal functions⁶ that are absolutely unnecessary in the discussion of the Hungarian word. Such cases are very frequent indeed. It is therefore rather unclear why polemic opinions are given without explanations, e.g. (in connection with Hung. *bögöly* ‘horsefly’): “As a base we can suppose an onomatopoeic verb **bö-* ~ **bü-* and fr[om] this a noun *bög* [...]. It is less likely that this also is a case of a nominal base **bö*, as claimed by Sevortjan [...], based on *bö*, *bög* ‘zehirli örümcek’ [= ‘poisonous spider’ – M. S.] (OttT); rather the latter data look like abbreviated written forms of *bög*” (I 168). It is out of the question that *bö* is an “abbreviated written form” (what does this even mean?) if only because no such “abbreviations” exist in Turkish at all. In addition, no argument is given in favour of disregarding the importance of the attested *bö* and, instead, supporting the onomatopoeic interpretation, actually based on nothing, not even a single onomatopoeic record or meaning in any Turkic language whatsoever.

csónak ‘boat’ (in older records with *-ln-*) < Slav. *čolnok* id. (II 1105). – The *-ln-* group is not only Old Hungarian; it is attested even today in a geographical name *Csolnok* (Stachowski M. 2009: 84).

ocsúdik ‘to awake’ (II 1214). – The authors rightly dismiss Tkcs. *ač-* ‘to open’ as an etymological source of Hung. *ocsúdik*. The semantic evolution: ‘to open’ > **‘to open one’s eyes’* > ‘to awake’ does not look natural and self-evident. Besides, the Hungarian word is only attested in the 16th century, and this fact has certainly affected the opinion of Hungarian etymologists: “Herleitung aus dem Atürk [= Alttürkischen – M. S.] ist kaum wahrsch[einlich]” (EWU II 1054). In this context, I would like to call attention to another fact: The verb for ‘to revive’ in Polish is *ocucić* (i.e. with a voiceless [-ućić]). It is only in South Poland, certain parts of which (e.g. Orawa) belonged to Hungary in the past, that this verb is pronounced *ocudzić* (with voiced [-uźić]), see SEJP 88. How do Polish dial. *ocudź-* and Hung. *ocsúd-* compare?

sisak ‘helmet’ (II 1234). – The newest (and actually the only professional) study devoted to this word family, Helimski 1997, is totally omitted here. Therefore, this section does not offer any correct or even acceptable solution.

⁴ See e.g. I 299: *mänzildin mänzilkä tüšä qopa barur ärdilär* ‘s ostanovki na ostanovku šli oni, ostanavlivajaš i podnimajaš’. Incidentally, it is not easy to say why most Russian quotations are left in Russian, whereas some others are translated into English, cf. I 300: Kirg. *tüş-* ‘to descend [...]’, Kzk. *tüs-* ‘spuskalsja, padal’.

⁵ See e.g. I 301: *yuxiya diš- ~ yuxiya düš-* ‘schlafen gehen’.

⁶ See e.g. I 302: “kak vspomogatel’nyj glagol oboznačael dejstvie, soveršaemoe stremitel’no, vnezapno”.

3.

A comparison of phonetic adaptation processes of Turkic loan-words in Hungarian with those of Pannonian Slavonic (= PannSlav.) may be viewed as a particularly tempting experiment. For instance, one may wonder if it was actually possible that the WOT **buγzai* ‘wheat’ was reflected by **būza* (I 186), not **būzā*, as well as that this **-a* changed into *-ā*. However, the Pannonian Slavonic materials confirm this evolutionary direction. As the PannSlav. **a* (i.e. **a/* = [ā]) yielded Old Hung. *ā* one is tempted to say that the Old Hungarian language tended to render the Slav. /a/ as *ā* rather, than as *ā̇* (Helimski 1988: 352sq. = 2000: 421sq.). The Old Hungarian word-final *-ā* developed into *-ā̇*, that is: PannSlav. *ā* > Old Hung. **ā* > (-) *ā̇* ~ *-ā̇* (Helimski l.c.). It is somewhat surprising that the auslaut **-ā̇* was both shortened and labialised. In addition, the Slavonic **a/* most probably was to some degree labial. Therefore I would rather posit a PannSlav. **ā̇* > Old Hung. **ā̇* > (-) *ā̇* ~ *-ā̇*. Hence, the WOT **buγzai*, too, could have first yielded a form like **būzā̇* or, preferably, **būzā̇̇* (in any case not **būza*, unlike the reconstruction in I 186) that would have later become **būzā̇̇̇*. In this case a Slavistic insight suggests a more detailed interpretation of a Turkic loan-word in Hungarian.

* *
*

There is no doubt that the book under review is an important work. It will be doubtless consulted again and again. It will inspire scholars and initiate discussions. On the other hand, however, it will not gain the status of unquestionable and internationally accepted authority because its Turcological part is more often than not one-sided and very individual and, thus, always needs comments and annotations.

Marek Stachowski
ul. Barska 1/4
PL – 30-307 Kraków
[stachowski.marek@gmail.com]

R e f e r e n c e s

- EWU = Benkő L. et al.: *Etymologisches Wörterbuch des Ungarischen*, Budapest, vol. I: 1993, vol. II: 1994.
SEJP = Boryś W.: *Słownik etymologiczny języka polskiego*, Kraków 2005.
TMEN = Doerfer G.: *Türkische und mongolische Elemente im Neupersischen*, Wiesbaden, vol. I–IV: 1963–1975.

- Benkő L. → EWU.
- Benzing J. 1944: Die angeblichen bolgartürkischen Lehnwörter im Ungarischen. – *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 98: 24–27.
- Boryś W. → SEJP.
- Doerfer G. → TMEN.
- Helimski E. 1988: Vengerskij jazyk kak istočnik dlja praslavjanskoj rekonstrukcii i rekonstrukcii slavjanskogo jazyka Pannonii. – Tolstoj N. I. (ed.): *Slavjanskoe jazykoznanie. X Meždunarodnyj s'ezd slavistov. Doklady sovetskoj delegacii*, Moskva: 347–368.
- 1997: Dva šiśá: Turkic šiś(ik) and Fennic hīsi in Russian. – *Studia Etymologica Cracoviensia* 2: 151–157.
- 2000: *Komparativistika, uralistika. Lekcii i stat'ii*, Moskva.
- Li Y.-S. 1997: On the origin of *baqa* ‘frog, toad; tortoise’. – *Central Asiatic Journal* 41: 250–269.
- Piowarczyk D. R. 2014: The Proto-Indo-European root for ‘apple’ and the problem of comparative reconstruction. – *Studia Etymologica Cracoviensia* 19: 161–167.
- Róna-Tas A. 1999: Chuvash and historical morphology. – *Acta Orientalia Acad. Sc. Hung.* 52: 1–15.
- Ščerbak A. M. 1970: *Sravnitel'naja fonetika tjurkskich jazykov*, Leningrad.
- Stachowski K. 2005: *Wampir* na rozdrożach. Etymologia wyrazu *upiór* ~ *wampir* w językach słowiańskich. – *Rocznik Slawistyczny* 55: 73–92.
- 2008: *Names of cereals in the Turkic languages*, Kraków.
- Stachowski M. 1998: Zwei alttürkische Konsonantenwechsel (*š* ~ *s*, *š* ~ *l*), die Runik und die Altaistik. – Laut J. P. / Ölmez M. (ed.): *Bahşi Ögdisi* [Festschrift K. Röhrborn], Freiburg – İstanbul: 391–399.
- 2003: On the problem of Oriental etymologies of Polish *ogar* and Hungarian *agár* ‘hound’. – *SEC* 8 (2003): 169–182.
- 2005: Uwagi o zapożyczeniach altajskich w języku prasłowiańskim i kwestie pokrewne. – Siemienieć-Golaś E. / Pomorska M. (ed.): *Turks and non-Turks. Studies on the history of linguistic and cultural contacts* [Festschrift S. Stachowski]: 437–454.
- 2009: Eugen Helimskis Materialien zur Erforschung der ältesten slawisch-ungarischen Sprachkontakte. – *Studia Etymologica Cracoviensia* 14: 35–107.